

## ABSTRACT OF THE DISCLOSURE

The presence of an installation on a data processing system may be detected by providing a signature that includes  $m$  files having paths associated therewith, respectively. A number  $n$  files on the data processing system are determined that  
5 match files in the signature and a files found ratio given by  $n/m$  is determined. A transformation is applied to the signature by replacing at least a portion of at least one of the paths with a new path. Then, a distance is determined between the  $n$  files on the data processing system and the  $m$  signature files. The distance corresponds to a sum of a number of path segments associated with the  $m$  signature files that cannot be  
10 matched to a corresponding path segment associated with files on the data processing system. The presence of the installation on the data processing system is determined based on the files found ratio and the distance.